

**Problem 1.**

True or False?

- (a) If A and B are disjoint events and both have positive probability, then A and B are independent.

True:  False:

Why:  $P(AB) = 0$  since  $A \cap B = \emptyset$  }  $\Rightarrow P(AB) \neq P(A)P(B)$   
 $P(A) \cdot P(B) > 0$

- (b) If  $0 < P(B) < 1$ , then  $P(A|B) + P(A|\bar{B}) = 1$

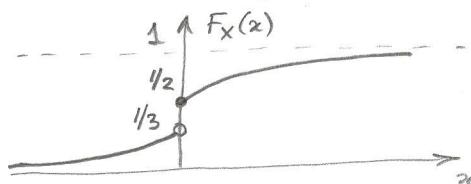
True:  False:

Why: Example: in tossing a fair die, let  $A = \{1\}$  and  $B = \{1, 2, 3\}$   
 $\Rightarrow \bar{B} = \{4, 5, 6\}$ ,  $P(A|B) = \frac{1}{3}$ ,  $P(A|\bar{B}) = 0$

- (c) There exists a random variable X such that the following function is the CDF of X.

$$F_X(x) = \begin{cases} \frac{e^x}{3}, & x < 0 \\ 1 - \frac{e^{-x}}{2}, & x \geq 0 \end{cases}$$

True:  False:



Why:  $F_X(x)$  is non-decreasing,  $\lim_{x \rightarrow -\infty} F_X(x) = 0$ ,  $\lim_{x \rightarrow \infty} F_X(x) = 1$ , and  $F_X(x)$  is right-continuous

- (d) It is possible to compute the correlation between  $X$  and  $Y = X - \mu$ , where  $\mu$  is unknown constant.

True:  False:

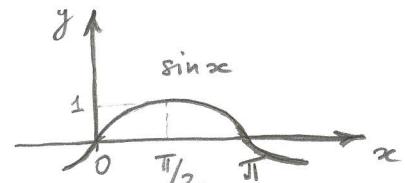
Why:  $\rho(X, X - \mu) = 1$

- (e) If  $X$  follows uniform distribution on  $[0, \pi]$  and  $f(x) = \sin(x)$ , then  $E[\sin(X)] \leq \sin(E[X])$ .

True:  False:

Why: Since  $g(x) = \sin x$  is a concave function on  $x \in [0, \pi]$ ,

$$E[\sin(X)] \leq \sin(E[X])$$



**Problem 2.** Suppose the 30% of computer owners use a Macintosh, 50% use Windows, and 20% use Linux. Suppose that 65% of the Mac users have succumbed to a computer virus, 82% of the Windows users get the virus, and 50% of the Linux users get the virus. We select a person at random and learn that her system was infected with the virus. What is the probability that she is a Windows user?

Solution : Let us introduce the following events :

$M$  = "the selected person is a Macintosh user"

$W$  = "the selected person is a Windows user"

$L$  = "the selected person is a Linux user"

Then  $\boxed{P(M) = 0.3, P(W) = 0.5, P(L) = 0.2}$

Let  $V$  be the event that the system is infected with the virus.

Then  $\boxed{P(V|M) = 0.65, P(V|W) = 0.82, P(V|L) = 0.5}$

Our goal is to find  $P(W|V)$ .

Using Bayes' theorem :

$$P(W|V) = \frac{P(V|W) \cdot P(W)}{P(V)}$$

Using the law of total probability :

$$P(V) = P(M) \cdot P(V|M) + P(W) \cdot P(V|W) + P(L) \cdot P(V|L)$$

Thus,

$$P(W|V) = \frac{P(V|W) P(W)}{P(M) P(V|M) + P(W) P(V|W) + P(L) P(V|L)} \approx 0.58$$

$$\boxed{P(W|V) \approx 0.58}$$

**Problem 3.**

Suppose that  $(X, Y)$  is uniformly distributed over the region defined by  $-1 \leq x \leq 1$  and  $0 \leq y \leq 1 - x^2$ .

(a) Find the marginal densities  $f_X(x)$  and  $f_Y(y)$  of  $X$  and  $Y$ .

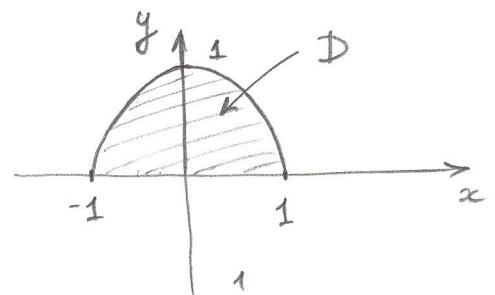
The joint distribution of  $X$  and  $Y$  is

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\text{Area}(D)} & \text{if } (x,y) \in D \\ 0 & \text{if } (x,y) \notin D \end{cases}$$

$$f_{X,Y}(x,y) = \begin{cases} \frac{3}{4} & \text{if } (x,y) \in D \\ 0 & \text{if } (x,y) \notin D \end{cases}$$

- $f_X(x) = \int_0^{1-x^2} f_{X,Y}(x,y) dy = \frac{3}{4} (1-x^2), \quad x \in [-1, 1]$

- $f_Y(y) = \int_{-\sqrt{1-y}}^{\sqrt{1-y}} f_{X,Y}(x,y) dx = \frac{3}{4} \cdot 2 \sqrt{1-y} = \frac{3}{2} \sqrt{1-y}, \quad y \in [0, 1]$



$$\begin{aligned} \text{Area} &= \int_{-1}^1 (1-x^2) dx = \\ &= \left( x - \frac{x^3}{3} \right) \Big|_{-1}^1 = \frac{4}{3} \end{aligned}$$

(b) Find the two conditional densities  $f_{X|Y}(x|y)$  and  $f_{Y|X}(y|x)$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{\frac{3}{4}}{\frac{3}{2} \sqrt{1-y}} = \frac{1}{2\sqrt{1-y}} \quad \begin{matrix} y \in [0, 1] \\ x \in [-\sqrt{1-y}, \sqrt{1-y}] \end{matrix}$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\frac{3}{4}}{\frac{3}{4} (1-x^2)} = \frac{1}{1-x^2} \quad \begin{matrix} x \in (-1, 1) \\ y \in [0, 1-x^2] \end{matrix}$$

$f_X(x) = \frac{3}{4} (1-x^2), \quad x \in [-1, 1]$
$f_Y(y) = \frac{3}{2} \sqrt{1-y}, \quad y \in [0, 1]$
$f_{X Y}(x y) = \frac{1}{2\sqrt{1-y}}, \quad y \in [0, 1], \quad x \in [-\sqrt{1-y}, \sqrt{1-y}]$
$f_{Y X}(y x) = \frac{1}{1-x^2}, \quad x \in (-1, 1), \quad y \in [0, 1-x^2]$

**Problem 4.** Suppose that  $X$  and  $Y$  are independent random variables with the standard deviations  $\sigma_x = 3$  and  $\sigma_y = 4$ . Let  $Z = Y - X$ . Find the covariance  $Cov(X, Z)$  and the correlation  $\rho(X, Z)$  of  $X$  and  $Z$ .

Solution  $\sigma_x = 3$ ,  $\sigma_y = 4$ ,  $Z = Y - X$ ,  $X$  and  $Y$  are independent

$$\begin{aligned} \bullet \quad Cov(X, Z) &= E[XZ] - E[X]E[Z] = E[X(Y-X)] - E[X] \cdot E[Y-X] \\ &= \underbrace{E[XY]}_{=} - E[X^2] - E[X]E[Y] + E[X]^2 = - (E[X^2] - E[X]^2) = - \sigma_x^2 \\ &= - 9 \end{aligned}$$

$E[X] \cdot E[Y]$  since  $X$  and  $Y$  are independent

$$\bullet \quad \rho(X, Z) = \frac{Cov(X, Z)}{\sigma_x \sigma_z} = - \frac{\sigma_x}{\sigma_z}$$

$$\sigma_z^2 = V[Z] = V[Y-X] = V[Y] + V[X] = \sigma_y^2 + \sigma_x^2$$

since  $X$  and  $Y$  are independent

$$\rho(X, Z) = - \frac{\sigma_x}{\sqrt{\sigma_x^2 + \sigma_y^2}} = - \frac{3}{5}$$

$$Cov(X, Z) = -9$$

$$\rho(X, Z) = - \frac{3}{5}$$

**Problem 5.** Suppose we have a computer program consisting of  $n=100$  pages of code. Let  $X_i$  be the number of errors on the  $i^{\text{th}}$  page of code. Suppose that the  $X_i$  are Poisson with mean  $\lambda=1$  and that they are independent. Let  $Y = \sum_{i=1}^n X_i$  be the total number of errors. Use the central limit theorem to approximate  $P(Y < 90)$ .

*Hint:* It may help to know that

1. if  $X \sim \text{Poisson}(\lambda)$ , then its variance is  $\sigma^2 = \lambda$ .
2.  $\Phi(-2) \approx 0.023$ ,  $\Phi(-1) \approx 0.159$ ,  $\Phi(0) = 0.5$

$$X_i \stackrel{iid}{\sim} \text{Poisson}(\lambda), \lambda = 1$$

$$Y = \sum_{i=1}^n X_i, n = 100$$

Solution: According to the CLT,  $\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ , where

$$\text{Thus, } \bar{X}_n \sim N\left(\lambda, \frac{\lambda^2}{n}\right) \quad \mu = E[X_i] = \lambda$$

$$Y = n \cdot \bar{X}_n \quad \sigma^2 = V[X_i] = \lambda$$

$$\begin{aligned} \Rightarrow P(Y < 90) &= P\left(n \bar{X}_n < 90\right) = P\left(\bar{X}_n < \frac{90}{n}\right) = \\ &= P\left(\underbrace{\frac{(\bar{X}_n - \lambda)\sqrt{n}}{\sqrt{\lambda}}}_{\sim N(0,1)} < \frac{\left(\frac{90}{n} - \lambda\right)\sqrt{n}}{\sqrt{\lambda}}\right) \approx \Phi\left(\frac{\left(\frac{90}{n} - \lambda\right)\sqrt{n}}{\sqrt{\lambda}}\right) = \\ &\quad = \Phi(-1) \approx 0.159 \end{aligned}$$

$$P(Y < 90) \approx 0.159$$